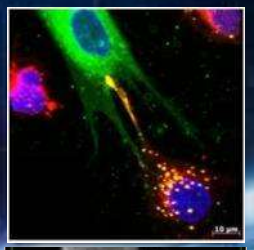
National Aeronautics and Space Administration

Human Research Program

The Twins Study: NASA's First Foray into 21st Century Omics Research

Grand Rounds
23 September 2014






Craig E. Kundrot, Ph.D.
Deputy Chief Scientist, HRP
SA2/NASA JSC

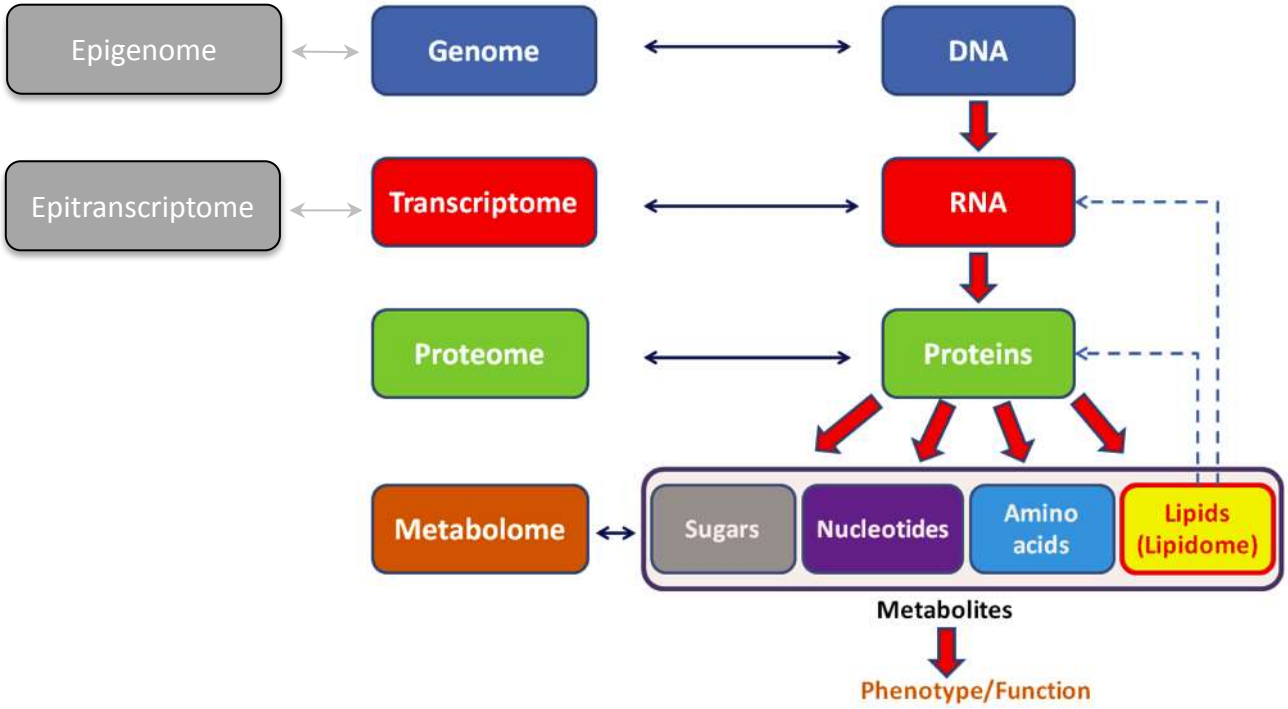
Outline

- What is “omics” and what can we learn from an omics investigation?
- What is the Twins Study?
- What issues is NASA grappling with as it undertakes omic research?

Omics

Omics: A neologism for the constellation of an organism's "-omic" information, which includes the genome itself (genomic), transcription products (transcriptomic), protein products (proteomic) and metabolic products (metabolomic).

medical-dictionary.thefreedictionary.com/omics



http://regionalextensioncenter.blogspot.com/2011_11_01_archive.html

Example: The "Snyderome"

Mike Snyder

Resource

Cell

Personal Omics Profiling Reveals Dynamic Molecular and Medical Phenotypes

Rui Chen,[1,11] George I. Mias,[1,11] Jennifer Li-Pook-Than,[1,11] Lihua Jiang,[1,11] Hugo Y.K. Lam,[1,12] Rong Chen,[2,12] Elana Miriami,[1] Konrad J. Karczewski,[1] Manoj Hariharan,[1] Frederick E. Dewey,[3] Yong Cheng,[1] Michael J. Clark,[1] Hogune Im,[1] Lukas Habegger,[4,7] Suganthi Balasubramanian,[4,7] Maeve O'Huallachain,[1] Joel T. Dudley,[5] Sara Hillenmeyer,[1] Rajini Haraksingh,[1] Donald Sharon,[1] Ghia Euskirchen,[1] Phil Lacroute,[1] Keith Bettinger,[1] Alan P. Boyle,[1] Maya Kasowski,[1] Fabian Grubert,[1] Scott Seki,[2] Marco Garcia,[2] Michelle Whirl-Carrillo,[1] Mercedes Gallardo,[6,10] Maria A. Blasco,[6] Peter L. Greenberg,[4] Phyllis Snyder,[1] Teri E. Klein,[1] Russ B. Altman,[1,5] Atul J. Butte,[2] Euan A. Ashley,[3] Mark Gerstein,[4,7,8] Kari C. Nadeau,[2] Hua Tang,[1] and Michael Snyder[1,*]
[1]Department of Genetics, Stanford University School of Medicine
[2]Division of Systems Medicine and Division of Immunology and Allergy, Department of Pediatrics
[3]Center for Inherited Cardiovascular Disease, Division of Cardiovascular Medicine
[4]Division of Hematology, Department of Medicine
[5]Department of Bioengineering
Stanford University, Stanford, CA 94305, USA
[6]Program in Computational Biology and Bioinformatics
[7]Department of Molecular Biophysics and Biochemistry
[8]Department of Computer Science
Yale University, New Haven, CT 06520, USA
[9]Telomeres and Telomerase Group, Molecular Oncology Program, Spanish National Cancer Centre (CNIO), Madrid E-28029, Spain
[10]Life Length, Madrid E-28003, Spain
[11]These authors contributed equally to this work
[12]Present address: Personalis, Palo Alto, CA 94301, USA
*Correspondence: mpsnyder@stanford.edu
DOI 10.1016/j.cell.2012.02.009

SUMMARY

Personalized medicine is expected to benefit from combining genomic information with regular monitoring of physiological states by multiple high-throughput methods. Here, we present an integrative personal omics profile (iPOP), an analysis that combines genomic, transcriptomic, proteomic, metabolomic, and autoantibody profiles from a single individual over a 14 month period. Our iPOP analysis revealed various medical risks, including type 2 diabetes. It also uncovered extensive, dynamic changes in diverse molecular components and biological pathways across healthy and diseased conditions. Extremely high-coverage genomic and transcriptomic data, which provide the basis of our iPOP, revealed extensive heteroallelic changes during healthy and diseased states and an unexpected RNA editing mechanism. This study demonstrates that longitudinal iPOP can be used to interpret healthy and diseased states by connecting genomic information with additional dynamic omics activity.

INTRODUCTION

Personalized medicine aims to assess medical risks, monitor, diagnose and treat patients according to their specific genetic composition and molecular phenotype. The advent of genome sequencing and the analysis of physiological states has proven to be powerful (Cancer Genome Atlas Research Network, 2011). However, its implementation for the analysis of otherwise healthy individuals for estimation of disease risk and medical interpretation is less clear. Much of the genome is difficult to interpret and many complex diseases, such as diabetes, neurological disorders and cancer, likely involve a large number of different genes and biological pathways (Ashley et al., 2010; Grayson et al., 2011; Li et al., 2011), as well as environmental contributors that can be difficult to assess. As such, the combination of genomic information along with a detailed molecular analysis of samples will be important for predicting, diagnosing and treating diseases as well as for understanding the onset, progression, and prevalence of disease states (Snyder et al., 2009).

Presently, healthy and diseased states are typically followed using a limited number of assays that analyze a small number of markers of distinct types. With the advancement of many new technologies, it is now possible to analyze upward of 10^9 molecular constituents. For example, DNA microarrays have allowed the subcategorization of lymphomas and gliomas

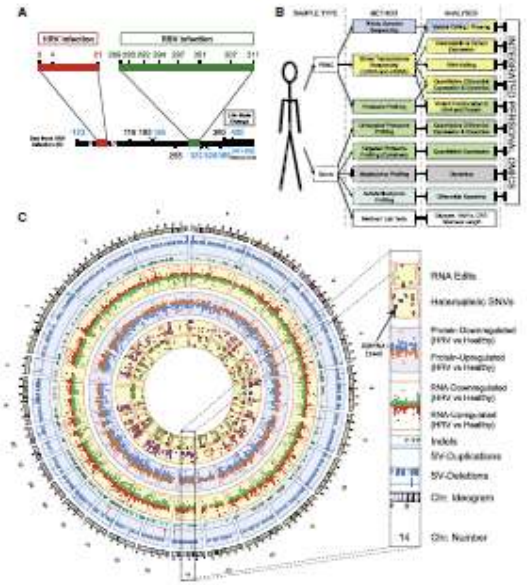


Figure 1. Summary of Study

(A) Time course summary. The subject was monitored for a total of 726 days, during which there were two infections (red bar, HRV; green bar, RSV). The black bar indicates the period when the subject: (1) increased exercise, (2) ingested 81 mg of acetylsalicylic acid and ibuprofen tablets each day (the latter only during the first 6 weeks of this period), and (3) substantially reduced sugar intake. Blue numbers indicate blood draw time points.

(B) iPOP experimental design indicating the tissues and analyses involved in this study.

(C) Circos (Krzywinski et al., 2009) plot summarizing iPOP. From outer to inner rings: chromosome ideogram; genomic data (pale blue ring), structural variants > 50 bp (deletions [blue tiles], duplications [red tiles]), indels (green triangles); transcriptomic data (yellow ring), expression ratio of HRV infection to healthy states; proteomic data (light purple ring), ratio of protein levels during HRV infection to healthy states; transcriptomic data (yellow ring), differential heteroallelic expression ratio of alternative allele to reference allele for missense and synonymous variants (purple dots) and candidate RNA missense and synonymous edits (red triangles, purple dots, orange triangles and green dots, respectively).

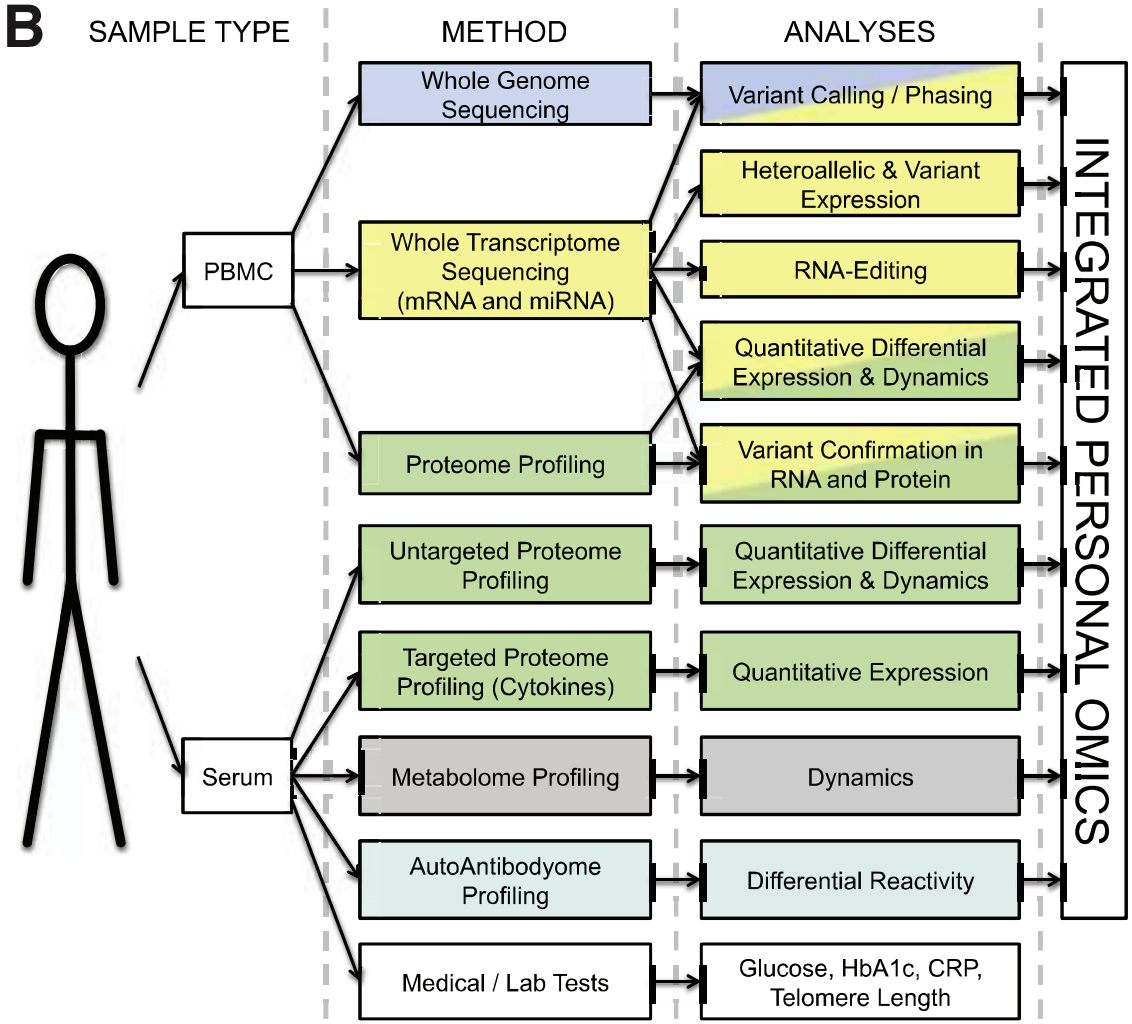
See also Figure S1.

WGS-Based Disease Risk Evaluation

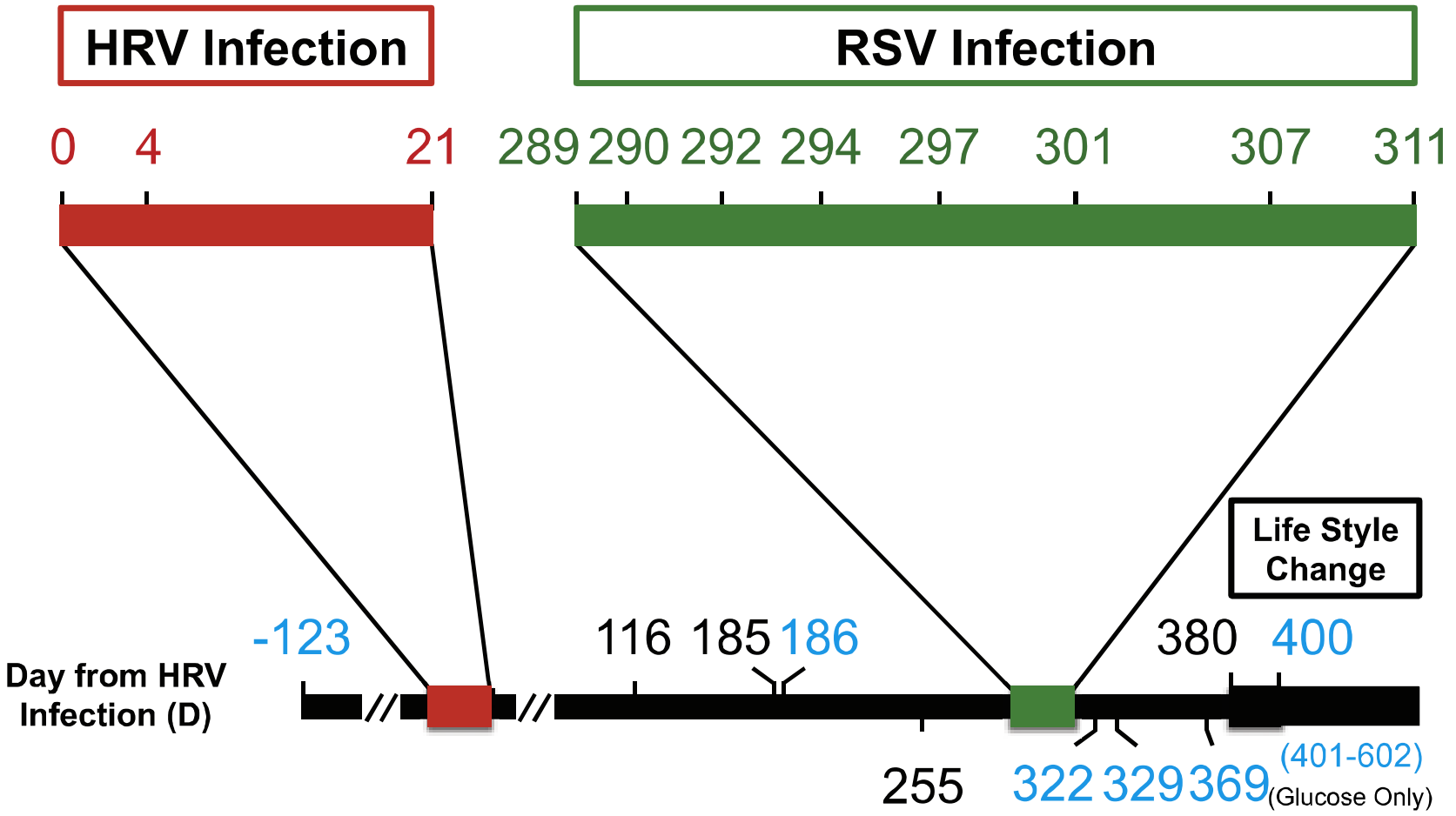
We identified variants likely to be associated with increased susceptibility to disease (Dewey et al., 2011). The list of high confidence SNVs and indels was analyzed for rare alleles (<5% of the major allele frequency in Europeans) and for changes in genes with known Mendelian disease phenotypes (data summarized in Table 2), revealing that 51 and 4 of the rare coding SNV and indels, respectively, in genes present in OMIM are predicted to lead to loss-of-function (Table S2A). This list of genes was further examined for medical relevance (Table S2A; example alleles are summarized in Figure 2A), and 11 were validated by Sanger sequencing. High interest genes include: (1) a mutation (E986K) in the *SERPINA7* gene previously known in the subject, (2) a damaging mutation in *TERT*, associated with acquired aplastic anemia (Yamaguchi et al., 2005), and (3) variants associated with hypertriglyceridemia and diabetes, such as *GCKR*

Celll (2012) 148:1293-1307.

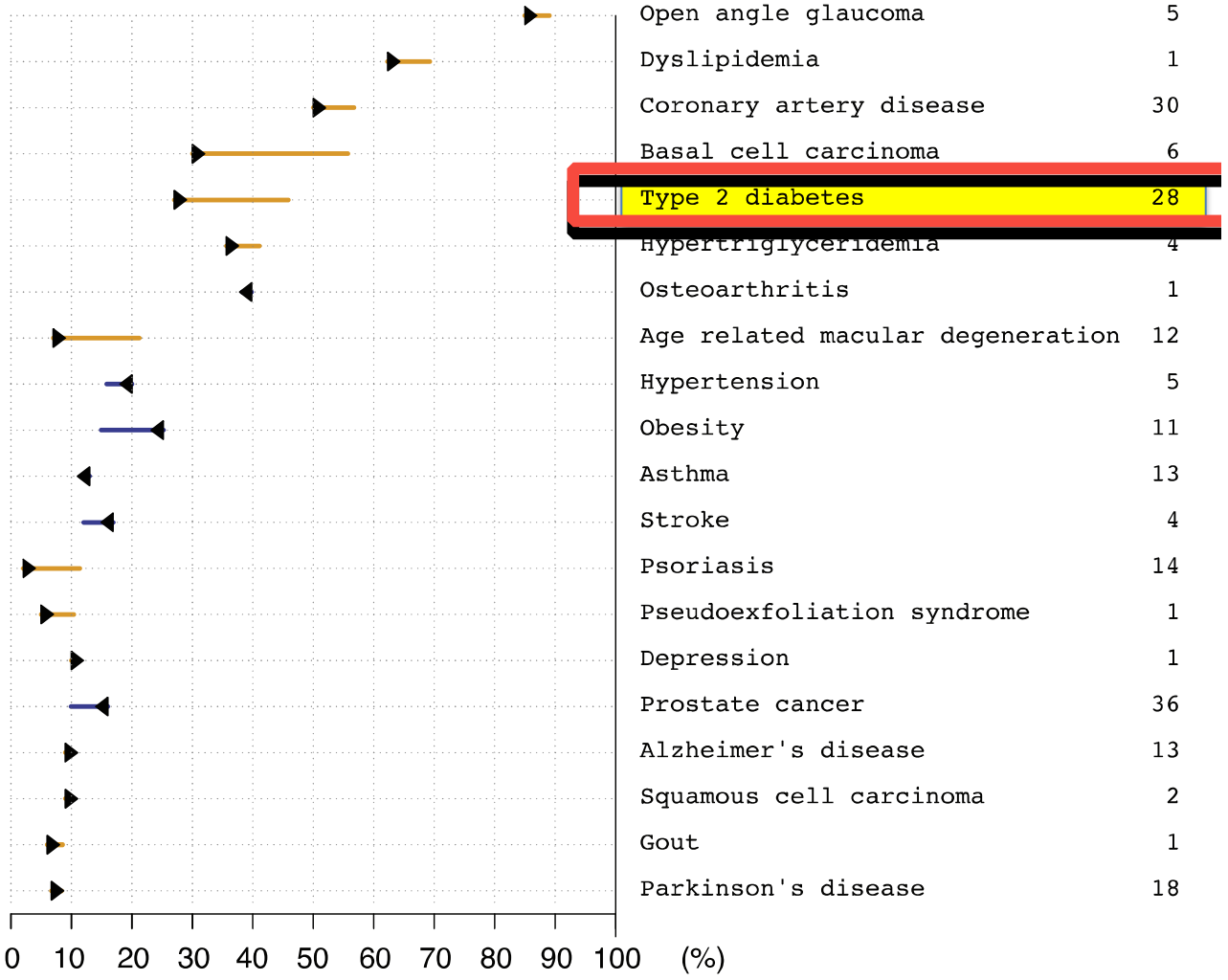
Analyses


B
SAMPLE TYPE
METHOD
ANALYSES
PBMC
Serum
Whole Genome Sequencing
Whole Transcriptome Sequencing (mRNA and miRNA)
Proteome Profiling
Untargeted Proteome Profiling
Targeted Proteome Profiling (Cytokines)
Metabolome Profiling
AutoAntibodyome Profiling
Medical / Lab Tests
Variant Calling / Phasing
Heteroallelic & Variant Expression
RNA-Editing
Quantitative Differential Expression & Dynamics
Variant Confirmation in RNA and Protein
Quantitative Differential Expression & Dynamics
Quantitative Expression
Dynamics
Differential Reactivity
Glucose, HbA1c, CRP, Telomere Length
INTEGRATED PERSONAL OMICS

Timeline

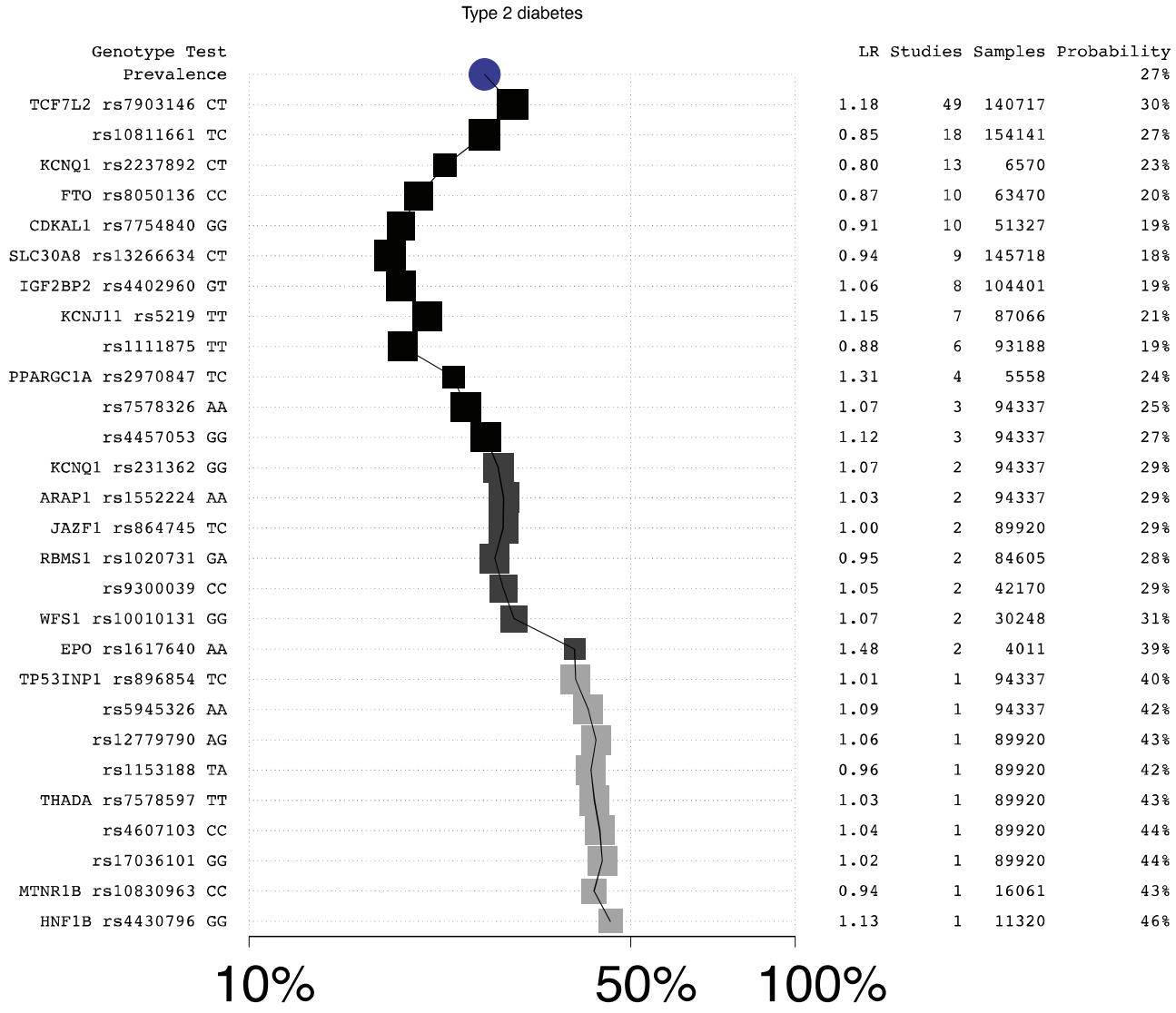

HRV Infection
RSV Infection
0
4
21
289
290
292
294
297
301
307
311
Life Style Change
Day from HRV Infection (D)
-123
116
185
186
380
400
255
322
329
369
(401-602)
(Glucose Only)

Genome: Quantitative Risk Estimates



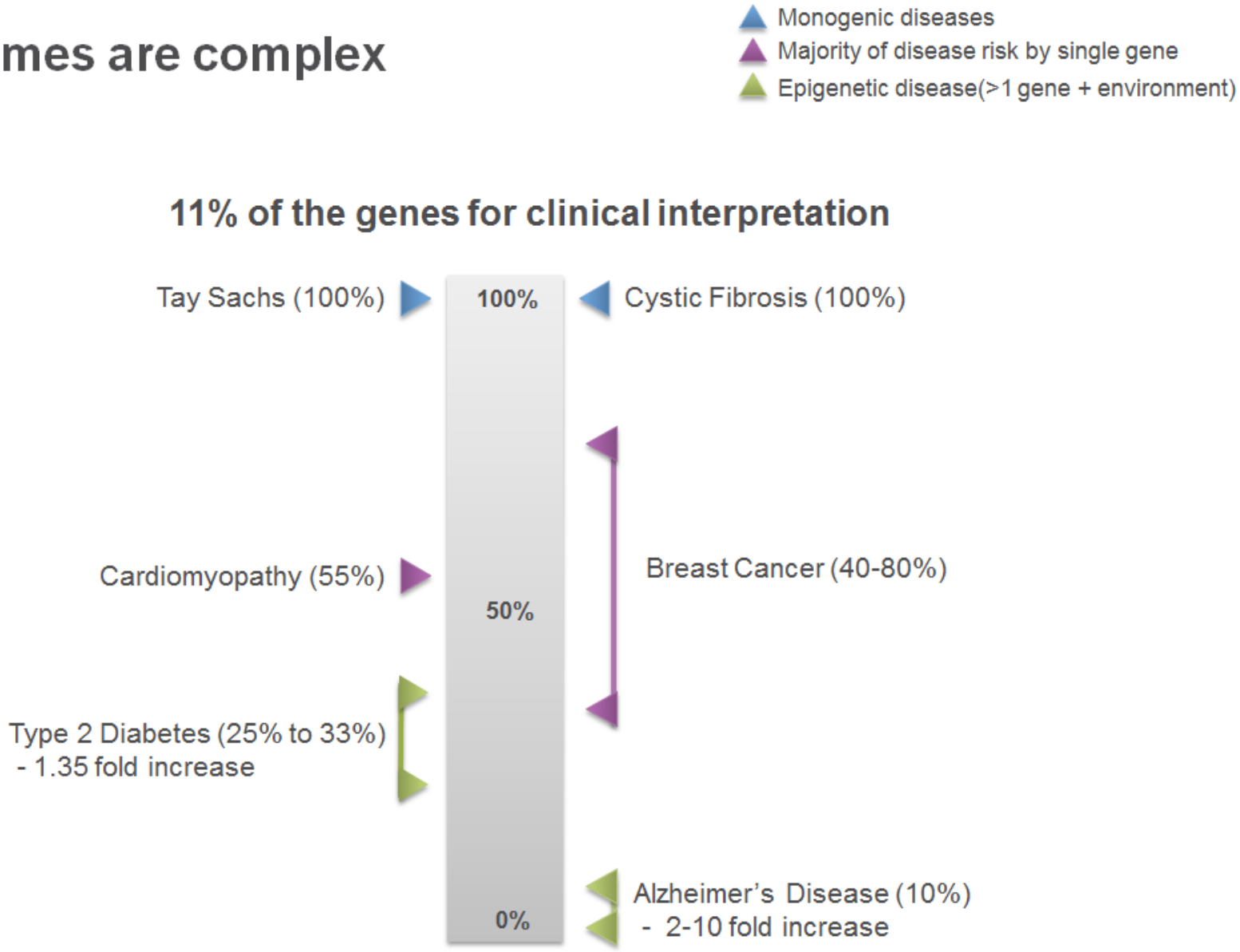
Decomposition of the Risk Estimate

C

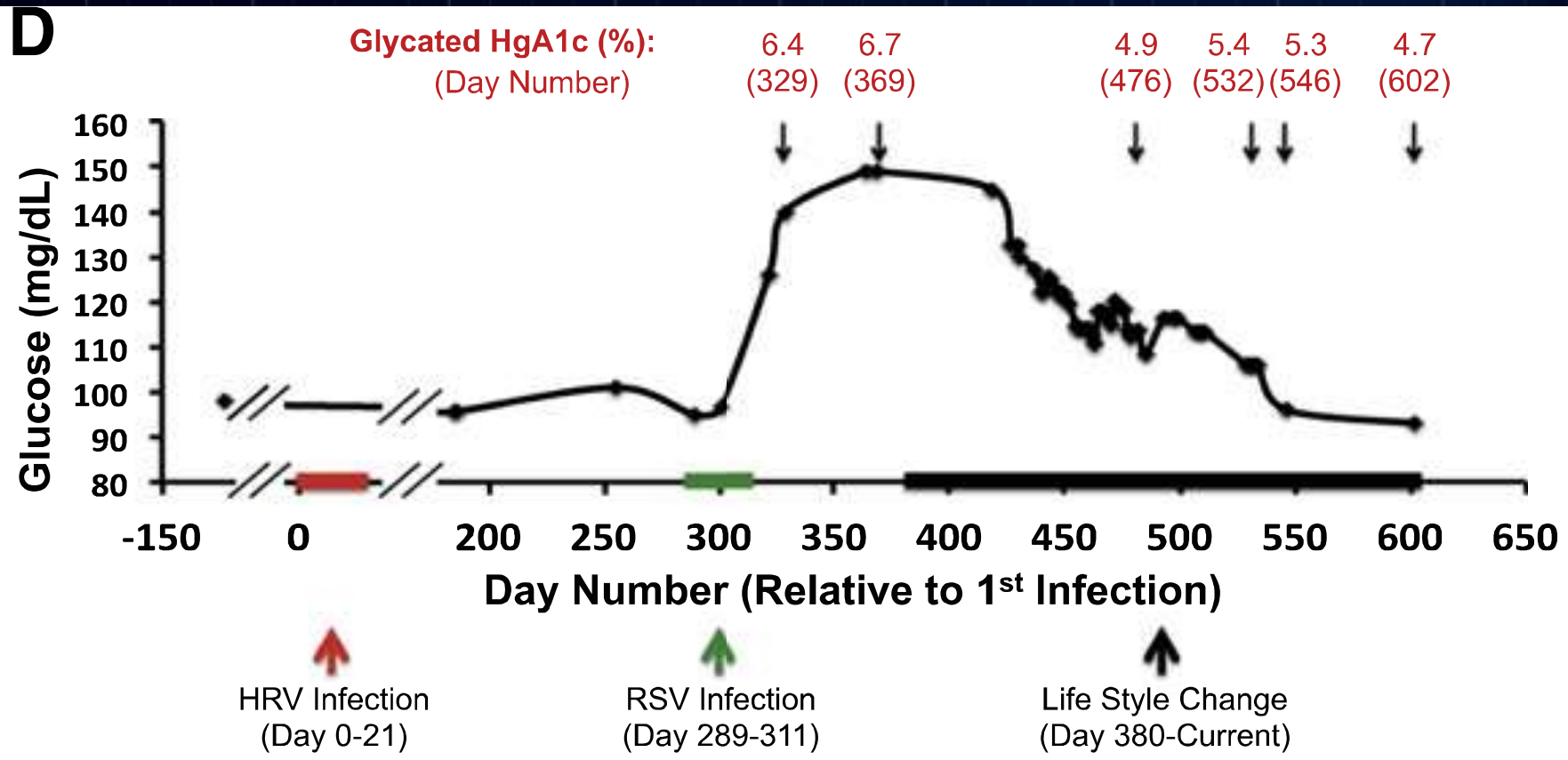


Wide Range of Predictive Power

Genomes are complex



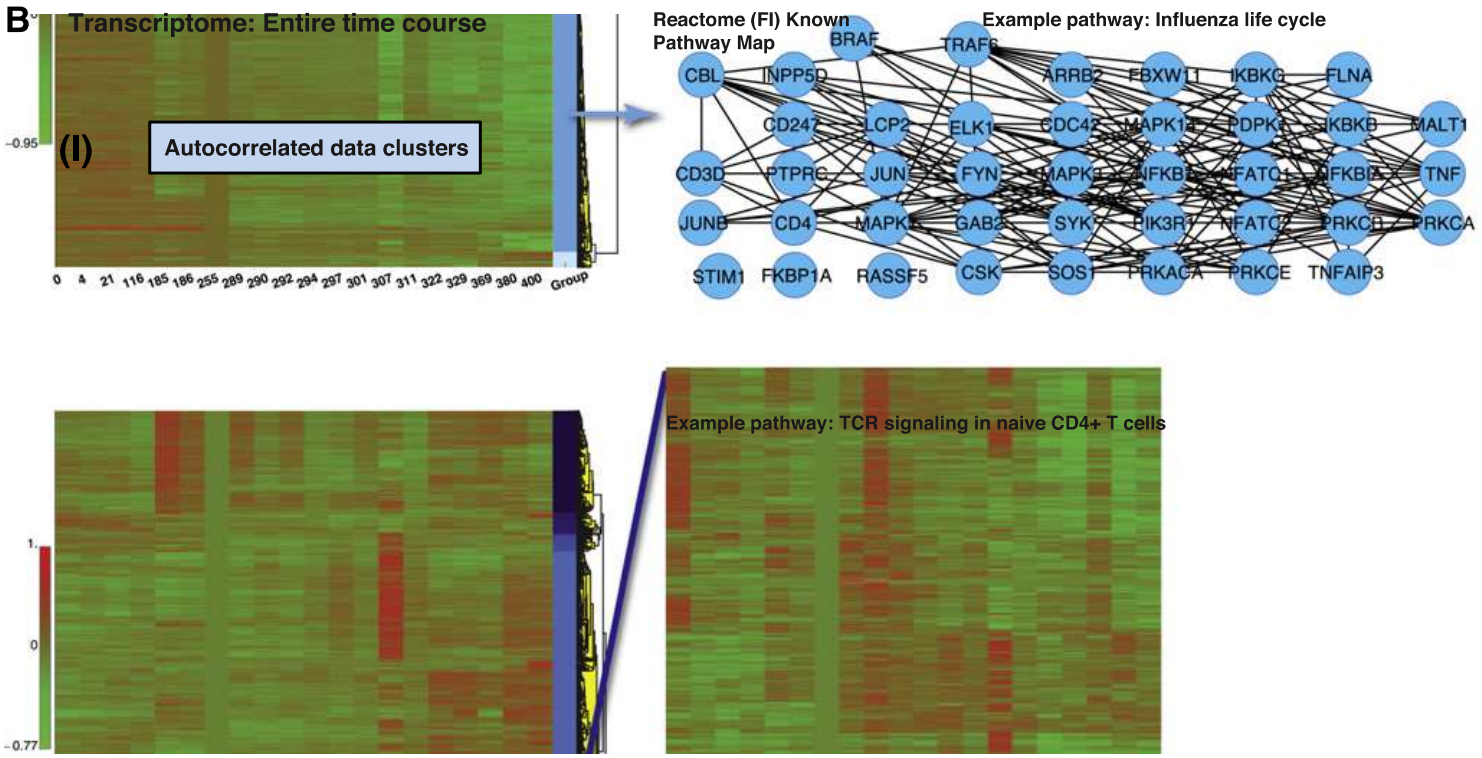
(Targeted) Metabolome: Glucose



"After a dramatic change in diet, exercise and ingestion of low doses of acetylsalicylic acid a gradual decrease in glucose (to ~93 mg/dl at day 602) and HbA1c levels to 4.7% was observed."

"These results indicate that a genome sequence can be used to estimate disease risk in a healthy individual, and by monitoring traits associated with that disease, disease markers can be detected and the phenotype treated."

Transcriptome: Unexpected Activations



"A large number of genes with a coexpression pattern common to both infections in the time course have yet to be implicated in known pathways and provide possible connections related to immune response."

Outline

- What is “omics” and what can we learn from an omics investigation?
- What is the Twins Study?
- What issues is NASA grappling with as it undertakes omic research?

The One-Year Mission

ISS Crew: Scott Kelly, Mikhail Kornienko Sign On For One-Year Mission

Posted: 11/26/2012 9:26 am EST Updated: 11/26/2012 9:40 am EST

Like 96 people like this. Sign Up to see what your friends like.

36 12 6 3 17

share tweet +1 email comment

GET SCIENCE ALERTS:

Enter email SIGN UP

REACT: Amazing Inspiring Funny Scary Hot Crazy Important Weird

FOLLOW: Video, Scott Kelly, International Space Station, Iss Crew, Iss Mission, Mikhail Kornienko, International Space Station, Science News

By: Tariq Malik
Published: 11/26/2012 08:12 AM EST on SPACE.com

A veteran NASA space commander and Russian cosmonaut have signed on for the ultimate space voyage: a yearlong trip on the International Space Station.

American astronaut Scott Kelly and Russian cosmonaut Mikhail Kornienko will launch on the one-year space station flight in spring 2015 and return to Earth in spring 2016, NASA officials announced today (Nov. 26). They will begin their mission training in early 2013.

The mission will help NASA understand how the human body adapts to extremely long space missions, such as voyages around the moon, to an asteroid and ultimately to Mars, NASA officials said.

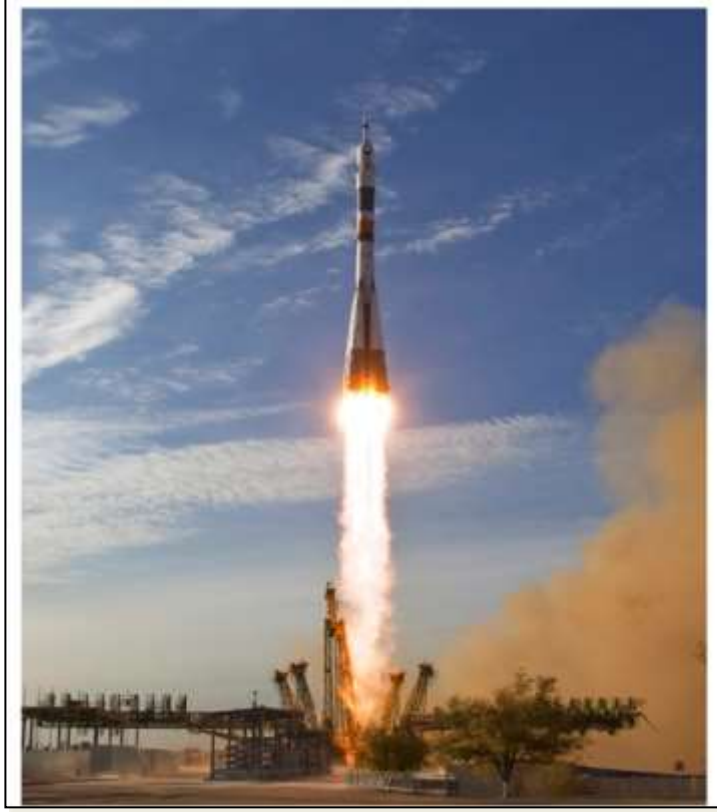
HOME > SCIENCE

Astronaut Scott Kelly Preparing for Unprecedented One Year in Space; Mission to Experiment on His Bone Mass, Vision, Immune System

By Latin Times Staff Writer, Dec 07, 2012 08:00 PM EST

0 Comments Like 0 Tweet 0 +1 0 Share Text Size

Tags: NASA, Space



The Twins Opportunity Originated at the Initiative of Scott & Mark Kelly

SCIENCE

30 COMMENTS

NASA will separate twin brothers for a year: one on Earth, one in space

By Carl Franzen on August 5, 2013 09:24 pm Email @carlfranzen

DON'T MISS STORIES FOLLOW THE VERGE Like 147k Follow 256K followers

THE LATEST HEADLINES

- New Oculus Rift cover shooter is so realistic, players warned not to lean on virtual objects
- Sea level and temperatures hit record highs last year, report says
- Microsoft launches Windows Phone web tool to let anyone create apps
- US government considers

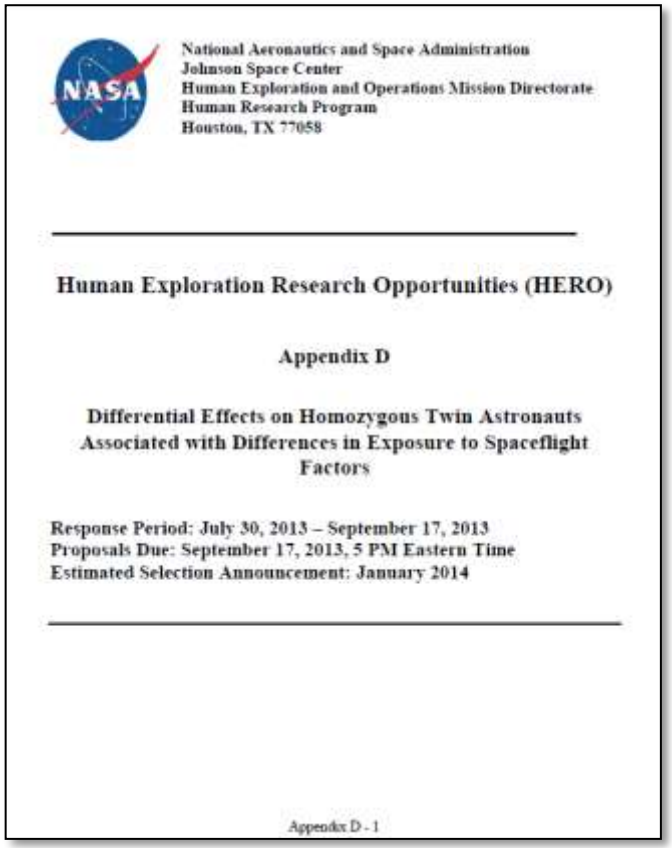
141

Of all the members of NASA's current crop of distinguished astronauts, only two have the unique distinction of being identical twin brothers. And now NASA is using an idea [illegible] the brothers, Mark Kelly and Scott Kelly, to perform a study that's been [illegible] up until now. Beginning in March 2015, the space agency will be [illegible] the biological states of both twin brothers over the span of a year, with a twist: Scott will be aboard the International Space Station for the duration of that period, while Mark, who retired from NASA back in 2011, will remain back here on Earth.

Tweet

THE VERGE

NRA Solicitation



National Aeronautics and Space Administration
Johnson Space Center
Human Exploration and Operations Mission Directorate
Human Research Program
Houston, TX 77058

Human Exploration Research Opportunities (HERO)

Appendix D

Differential Effects on Homozygous Twin Astronauts Associated with Differences in Exposure to Spaceflight Factors

Response Period: July 30, 2013 – September 17, 2013
Proposals Due: September 17, 2013, 5 PM Eastern Time
Estimated Selection Announcement: January 2014

Appendix D - 1

"To capitalize on this unique opportunity,

NASA's Human Research Program (HRP) and the National Space Biomedical Research Institute (NSBRI) are initiating

a *pilot demonstration project focused on the use of integrated human -omic analyses* to

better understand the biomolecular responses to

the physical,
physiological, and
environmental stressors associated with

spaceflight."

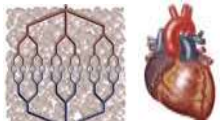
Selections

NASA Funded 10 Research Proposals in Response to its "Twins" Solicitation

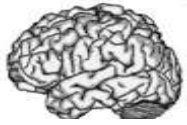
- 2 Subjects
  - Scott Kelly
  - Mark Kelly
- 10 Selections

1. Susan Bailey, Colorado State University, Differential effects on telomeres and telomerase in twin astronauts associated with spaceflight
2. Andrew Feinberg, Johns Hopkins University School of Medicine, Comprehensive whole genome analysis of differential epigenetic effects of space travel on monozygotic twins
3. Christopher Mason, Weill Medical College of Cornell University, The Landscape of DNA and RNA Methylation Before, During, and After Human Space Travel
4. Scott Smith, NASA Johnson Space Center, Biochemical Profile: Homozygous Twin control for a 12 month Space Flight Exposure
5. Emmanuel Mignot, Stanford University School of Medicine, HERO Twin Astronaut Study Consortium (TASC): Immunome Changes in Space
6. Fred Turek, Northwestern University, HERO Twin Astronaut Study Consortium (TASC) Project: Metagenomic Sequencing of the Bacteriome in GI Tract of Twin Astronauts
7. Stuart Lee, Wyle Laboratories, Metabolomic And Genomic Markers Of Atherosclerosis As Related To Oxidative Stress, Inflammation, And Vascular Function In Twin Astronauts
8. Brinda Rana, University of California, Proteomic Assessment of Fluid Shifts and Association with Visual Impairment and Intracranial Pressure in Twin Astronauts
9. Mathias Basner, University of Pennsylvania School of Medicine, HERO Twin Astronaut Study Consortium (TASC) Project: Cognition on Monozygotic Twin on Earth
10. Michael Snyder, Stanford University, HERO Twin Astronaut Study Consortium (TASC) Project: Longitudinal integrated multi-omics analysis of the biomolecular effects of space travel

http://www.nasa.gov/content/nasa-selects-10-proposals-to-explore-genetic-aspects-of-spaceflight/

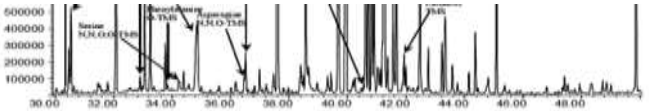


Vasculature
Lee



Cognition
Basner

Microbiome
Turek



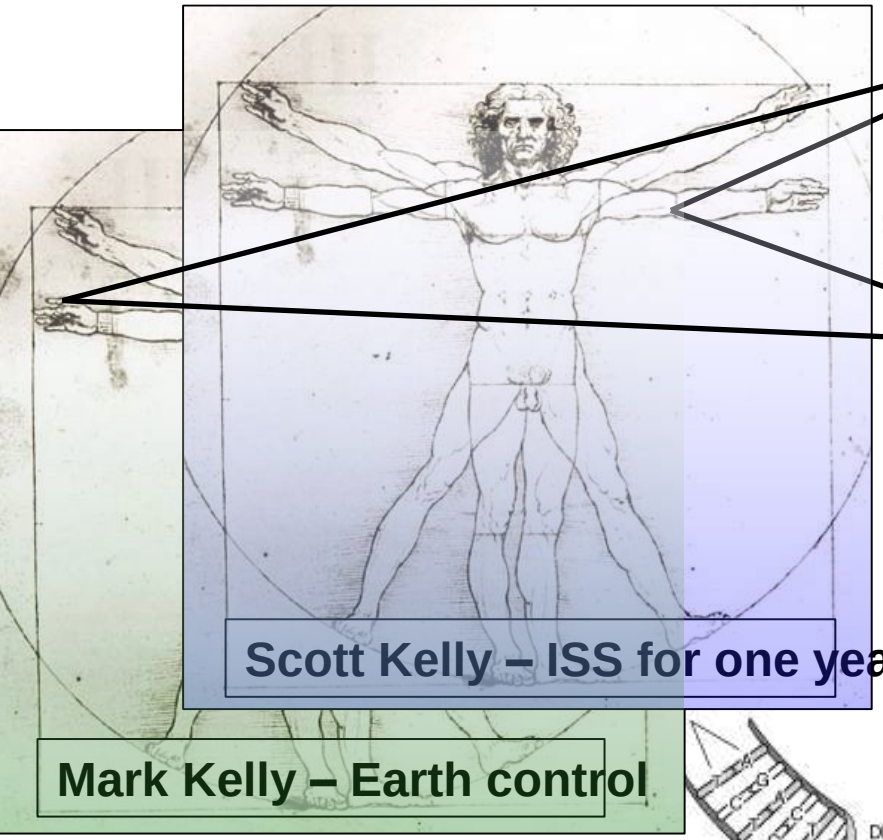
Targeted and Global Metabolomics
Lee/Rana, Mignot/Snyder & Smith

Cytokines
Mignot

Telomere Length
Bailey

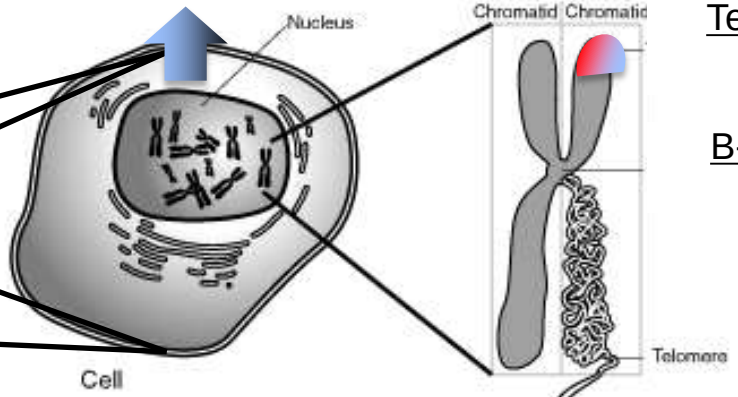
B-cells / T-cells
Mignot

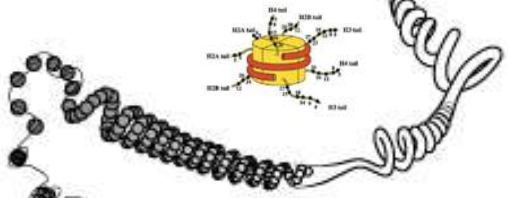
Antibodies
Mignot/Snyder



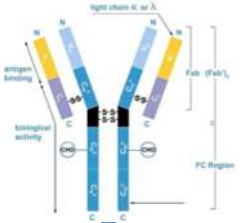
Scott Kelly – ISS for one year

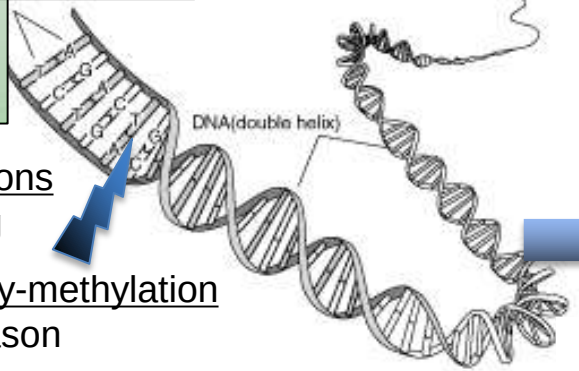
Mark Kelly – Earth control





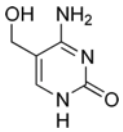
Chromatin
Feinberg

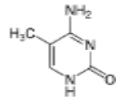




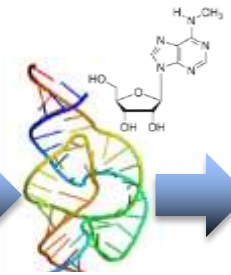
DNA Mutations
Feinberg

DNA Hydroxy-methylation
Mason

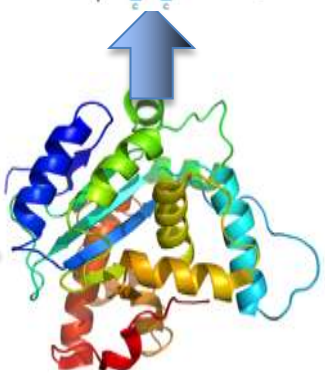




DNA Methylation
Feinberg & Mason

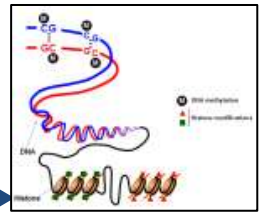


large/small RNA
& RNA Methylation
Mason



Proteomics
Lee/Rana

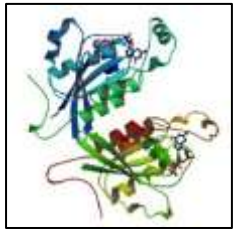
Buccal



Epigenome

DNA

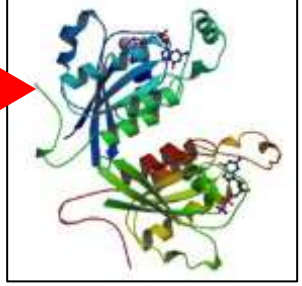
Urine



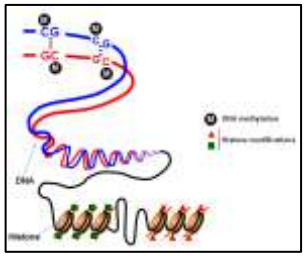
Proteins

Metabolites

Blood



Proteins

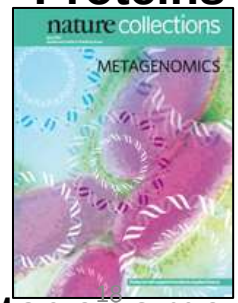


Epigenome

RNA

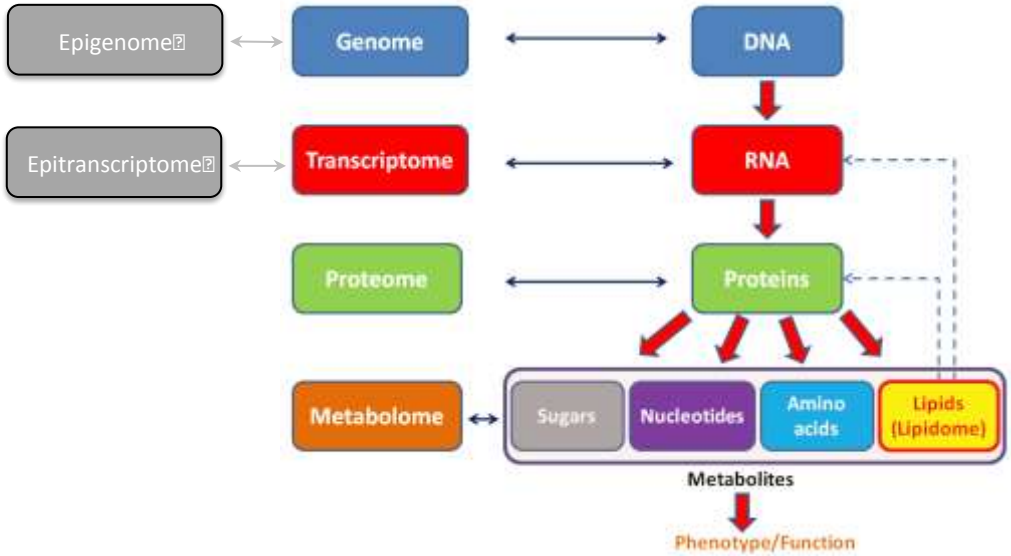
DNA

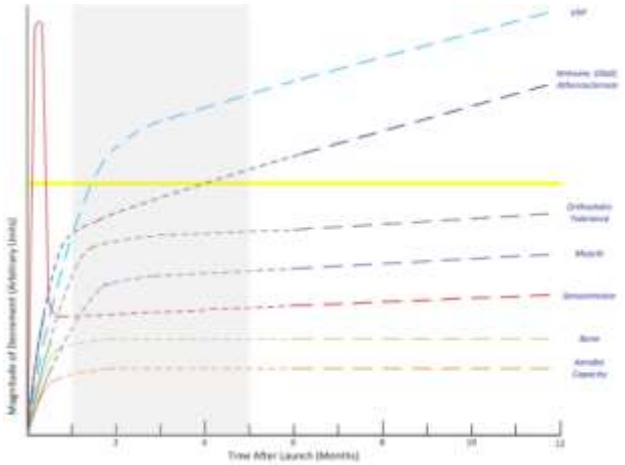
Stool

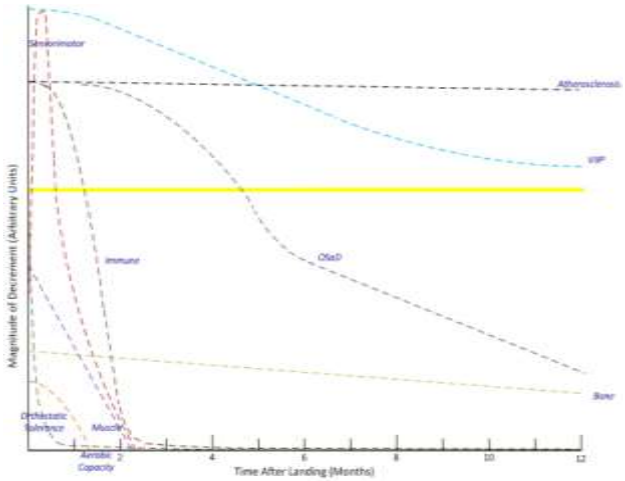


Metagenome

Measuring the Temporal Response to Space Flight







- 2 major sample collections pre-flight
- 10 major sample collections in-flight
- 2 major sample collections post-flight

- 6 major sample collections ground

Notional Time Courses

Outline

- What is “omics” and what can we learn from an omics investigation?
- What is the Twins Study?
- What issues is NASA grappling with as it undertakes omic research?

Issues Associated with Omic Research

- Protect the Research Subject
- Medical care
- Occupational health
- Employment activity

Protecting the Human Subject

- Interim policy on genetic research JID 1800.4
  - Applies to the NASA Flight IRB
  - "For purposes of this policy, the term 'genetic analysis' includes research involving human DNA, RNA, chromosomes, proteins, or metabolites that detects genotypes, mutations, or chromosomal changes. It excludes the analysis and collection of bio-specimens that will not be submitted to genetic analysis."

- Changes to the Informed Consent Form

Risks of Genetic Research

- The primary risks involved in genetic research are risks of social and psychological harm, rather than risks of physical injury

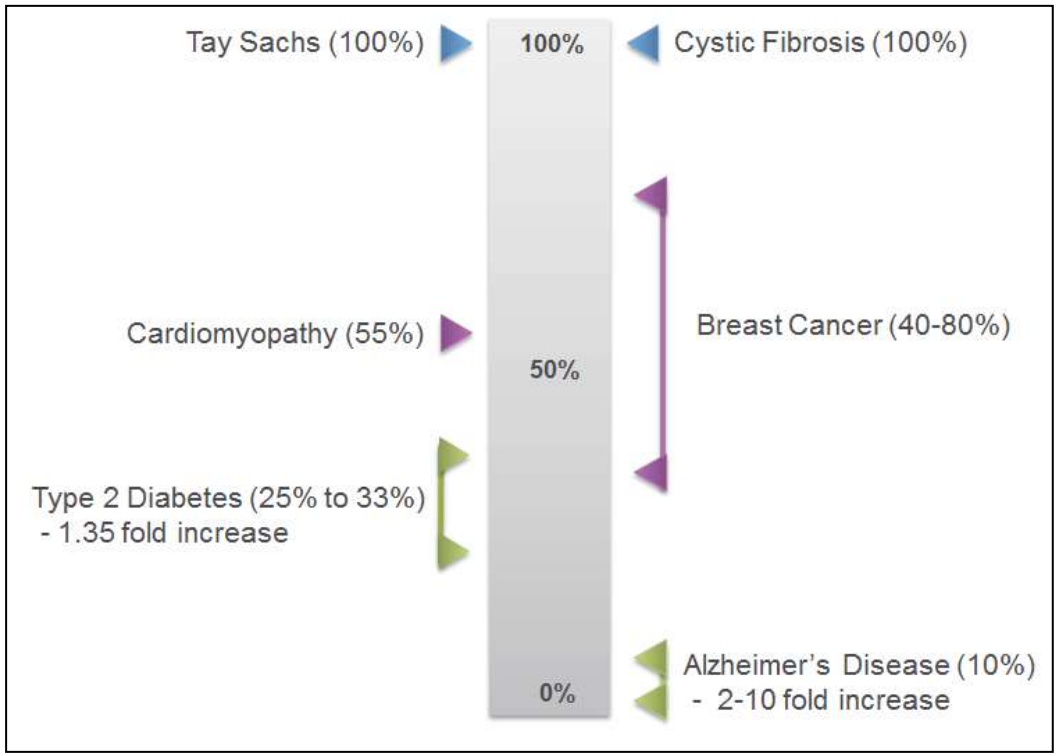
- Genetic studies that generate information about subjects' personal health risks
  - Could provoke anxiety and confusion
  - Damage familial relationships
  - Uncover unwanted information about heritage, ancestry, and family relationships

II.C. Additional Informed Consent Requirements

1. Any study involving genetic data shall provide test subjects with genetic counseling as appropriate to the study objectives and when requested by the Flight IRB.

Incidental Findings

- 56 genes might lead to medically actionable results
  - American College of Medical Genetics and Genomics (ACMG) 2013 – (http://goo.gl/C888BY)



II.C. Additional Informed Consent Requirements

1. Any study involving genetic data shall provide test subjects with genetic counseling as appropriate to the study objectives and when requested by the Flight IRB.

Sharing Information with the Research Subject

- Will the subject have the option to receive individual genome sequence data?
- Will the investigators interpret the results of the genome sequence and will that result be disclosed to the research subject?
- If the genome data are given to research subject will he/she have the option to decline to receive all or part of the results?

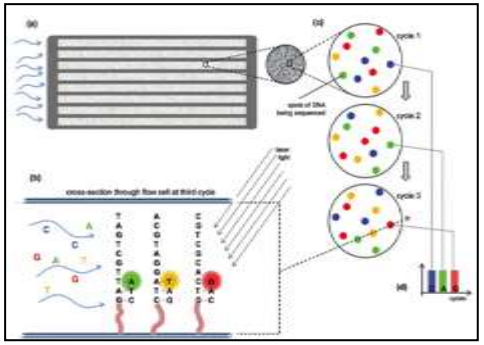
  (Right Not to Know)
- If there are medically actionable results will the investigators provide expert counseling or referral?

II.C. Additional Informed Consent Requirements

1. Any study involving genetic data shall provide test subjects with genetic counseling as appropriate to the study objectives and when requested by the Flight IRB.

Concern: Identifiable Data

- There are several kinds of gene sequencing
  - Whole Genome
  - RNA-seq
  - ChIP-seq
  - Methyl-seq



- All generate identifiable data
  - Deidentified sequences can be re-identified

Confidentiality and Privacy

- Individual genome sequences are unique and therefore uniquely identifiable

- Genome sequence placed in the public domain, may enable others to infer health information about the individual and his/her relatives

II.B.2. Use of genetic data

a) Data **shall not be disseminated** beyond the immediate control of the individual investigators documented in the protocol approved by the Flight IRB.

b) Genetic data **shall not be data-mined or cross-referenced with other databases** of any kind unless approved in advance by the Flight IRB.

c) Investigators **shall not attempt to identify individual participants** within de-identified data sets or pooled specimens, or to otherwise "reverse engineer" or "disassemble" data sets for bio-specimens involving NASA research subjects.

II.B.3. Security and Storage of data

a) Genetic data shall be encrypted and stored on secure servers. All IT systems used to store, process, or analyze genetic data shall comply with NASA's IT security standards for systems containing Privacy Act-protected information. In addition, no genetic data may be stored on mobile devices such as tablets, smart phones, or on removable media.

b) Genetic data stored on laptops shall be limited to the minimum amount required at any one time for current research purposes.

c) Once a study is completed, attributable data shall be archived in a secure manner by the investigator. **Investigators may be required to archive original study data at NASA or elsewhere at NASA's direction, and to destroy all copies of the original study data after the study is complete.**

II.B.4. Release of data

a) **No genetic sequence data may be posted online or otherwise published or made public**.

b) The IRB may waive this prohibition for the release of limited sequence data that is non-attributable. **The IRB must approve such a limited release in advance. The informed consent of the affected research subjects will be sought prior to such release.**

c) The **privacy** of genetic information will be **protected** to the full extent of the law, **including after the death of the subject** to avoid the unwarranted invasion of personal privacy of surviving family members.

The General Levels of Confidentiality and Privacy

1. The study may retain genome sequences and not allow any data sharing to any third party
2. The study may share with qualified third parties conducting related research
3. The study may share the data using a secure server like the NIH dbGAP
4. The study may deposit the data in a publically accessible database

II.B.3. Security and Storage of data
c) Once a study is completed, attributable data shall be archived in a secure manner by the investigator. Investigators may be required to archive original study data at NASA or elsewhere at NASA's direction, and to destroy all copies of the original study data after the study is complete.

II.B.4. Release of data
a) No genetic sequence data may be posted online or otherwise published or made public.
b) The IRB may waive this prohibition for the release of limited sequence data that is non-attributable. The IRB must approve such a limited release in advance. The informed consent of the affected research subjects will be sought prior to such release.
c) The privacy of genetic information will be protected to the full extent of the law, including after the death of the subject to avoid the unwarranted invasion of personal privacy of surviving family members.

II.C.3. **Subjects have the right to review presentation slides prior to public presentation and to review manuscript drafts and final publications before public release. Research subjects have the right to have their identifiable information removed from the presentation or publication.**

Insurance and Employment

- What assurances can be made about health insurance, disability insurance, life insurance, and employment?

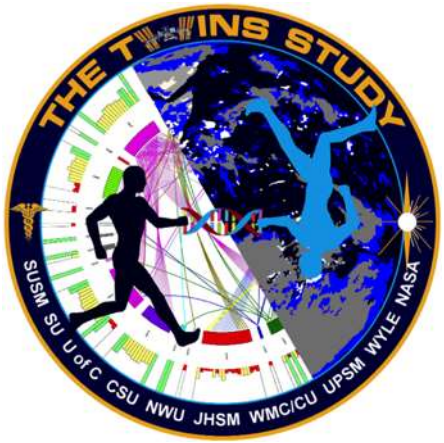
From the Informed Consent Form (ICF)

12.9 Confidentiality and release of protected health information for genetic research studies

- We will not convey your individual research results from this study to your medical record.
- We will not give your results to anyone else including your doctors. If we find something in your research testing that we believe can be used to directly help you with medical decisions, we will give that information to you.

Conclusion

- The Twins Study (Scott and Mark Kelly) is NASA's first foray into 21st-century omics research
  - Built around Scott Kelly's one year mission
- The Twins Study will examine
  - Genome and epigenome
  - Transcriptome and epitranscriptome
  - Proteome
  - Metabolome
  - Microbiome
  - Physiology
  - Cognition
- NASA is addressing
  - Protections for research subjects
    - Interim Genetic Research Policy JID 1800.4
    - Agency-level policy expected by summer 2016
  - Use of data in medical care and occupational medicine
  - Use of data in mission planning



Acknowledgements

The Twins Study Investigator Team

John Charles

Graham Scott

Bill Paloski

Mark Shelhamer

Jeff Sutton

Thank you